

JPA 09-114931 which corresponds to USP 6,137,905

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 09-114931

(43)Date of publication of application : 02.05.1997

(51)Int. Cl.

G06K 9/62

G06K 9/62

(21)Application number : 07-269826

(71)Applicant : CANON INC

(22)Date of filing : 18.10.1995

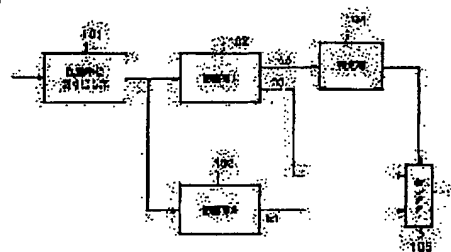
(72)Inventor : TODA YUKARI

(54) DEVICE AND METHOD FOR RECOGNIZING CHARACTER

(57)Abstract

PROBLEM TO BE SOLVED: To enable the character recognition of characters used for respective plural languages and to accelerate the processing speed of character recognition.

SOLUTION: This device is provided with recognizing parts 1 (102) and 2 (103) for recognizing the characters respectively used for Japanese and English, and concerning image data, the recognition is executed by the respective recognizing parts. The recognition rates of the respective recognized results are acquired, the respective acquired recognition rates are judged by a judging part 104 and based on that judgement, any one of the recognized results is selected by a selector 105. Then, the selected recognized result is outputted.



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

BEST AVAILABLE COPY

JPA09-114931 which corresponds to USP 6,137,905

(19) 日本国特許庁 (JP)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平9-114931

(43) 公開日 平成9年(1997)5月2日

(51) Int.Cl. ⁶	識別記号	庁内整理番号	FI	技術表示箇所
G 0 6 K 9/62	6 1 0	9061-5H	G 0 6 K 9/62	6 1 0 D
	6 2 0	9061-5H		6 2 0 B

審査請求 未請求 請求項の数28 OL (全 17 頁)

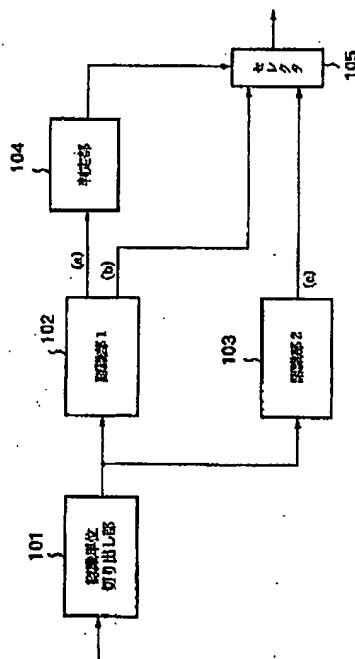
(21) 出願番号	特願平7-289828	(71) 出願人	000001007 キヤノン株式会社 東京都大田区下丸子3丁目30番2号
(22) 出願日	平成7年(1995)10月18日	(72) 発明者	戸田 ゆかり 東京都大田区下丸子3丁目30番2号 キヤノン株式会社内
		(74) 代理人	弁理士 大塚 康徳 (外1名)

(54) 【発明の名称】 文字認識装置及びその方法

(57) 【要約】

【課題】 複数言語のそれぞれに用いられる文字の文字認識を可能とし、かつ文字認識の処理速度を向上する文字認識装置及びその方法を提供する。

【解決手段】 日本語と英語のそれぞれで用いられる文字を認識する認識部1(102)と認識部2(103)を設け、画像データに関し、各認識部による認識を実行する。夫々の認識結果の認識率を獲得し、獲得される各認識率を判定部104で判定しその判定に基づいて、認識結果のどちらかをセレクト105で選択する。そして、選択される認識結果を出力する。



【特許請求の範囲】

【請求項1】 原稿を読み取って得られる画像データの文字認識を行う文字認識装置であって、
 複数種類の言語のそれぞれで用いられる文字を認識する複数の認識手段と、
 前記画像データに関し、前記複数の認識手段による認識を実行する実行手段と、
 前記認識手段による認識で得られる夫々の認識結果の認識率を獲得する獲得手段と、
 前記獲得手段によって獲得される前記複数の認識率に基づいて、複数の認識結果のいずれか一つを選択する選択手段と、
 前記選択手段によって選択される認識結果を出力する出力手段とを備えることを特徴とする文字認識装置。
 【請求項2】 前記選択手段は、前記複数の認識結果のそれぞれの認識率を比較する比較手段を更に備え、
 前記比較手段の比較の結果、認識率が一番高い認識結果を選択することを特徴とする請求項1に記載の文字認識装置。
 【請求項3】 前記選択手段は、前記認識結果のそれぞれの認識率を所定の閾値とで比較する比較手段とを更に備え、
 前記比較手段の比較の結果、前記所定の閾値より大きい認識結果を選択することを特徴とする請求項1に記載の文字認識装置。
 【請求項4】 前記原稿を光学的に読み取る読取手段と前記読取手段によって読み取られる画像データの属性に基づいて、前記画像データを属性毎の領域に分離する分離手段とを更に備えることを特徴とする請求項1に記載の文字認識装置。
 【請求項5】 前記認識手段は、前記分離手段によって分離される領域の属性が文字である領域を画像データとして認識することを特徴とする請求項4に記載の文字認識装置。
 【請求項6】 前記領域を所定の単位で分割する分割手段とを更に備えることを特徴とする請求項4に記載の文字認識装置。
 【請求項7】 前記所定の単位は、少なくとも行、文字を含むことを特徴とする請求項6に記載の文字認識装置。
 【請求項8】 原稿を読み取って得られる画像データの文字認識を行う文字認識装置であって、
 複数種類の言語のそれぞれで用いられる文字を認識する複数の認識手段と、
 前記複数の認識手段の一つを選択し、選択される認識手段による前記画像データの認識を実行する実行手段と、
 前記認識手段の認識によって得られる認識結果の認識率を獲得する獲得手段と、
 前記獲得手段によって獲得された前記認識率と所定の閾値とを比較する比較手段と、

前記比較手段の比較の結果、前記認識率が前記所定の閾値よりも大きい場合、前記認識結果を出力する出力手段と、
 前記出力手段による出力結果が得られるまで、前記認識手段の選択を順次切り換えて前記実行手段の実行がなされるよう制御する制御手段とを備えることを特徴とする文字認識装置。

【請求項9】 前記認識手段の使用順序を設定する設定手段を更に備え、

10 前記切り換えは、前記設定手段で設定される順序に従って切り換えることを特徴とする請求項8に記載の文字認識装置。

【請求項10】 前記原稿を光学的に読み取る読取手段と前記読取手段によって読み取られる画像データの属性に基づいて、前記画像データを属性毎の領域に分離する分離手段とを更に備えることを特徴とする請求項8に記載の文字認識装置。

【請求項11】 前記認識手段は、前記分離手段によって分離される領域の属性が文字である領域を画像データとして認識することを特徴とする請求項10に記載の文字認識装置。

【請求項12】 前記領域を所定の単位で分割する分割手段とを更に備えることを特徴とする請求項10に記載の文字認識装置。

【請求項13】 前記所定の単位は、少なくとも行、文字を含むことを特徴とする請求項12に記載の文字認識装置。

【請求項14】 原稿を読み取って得られる画像データの文字認識を行う文字認識方法であって、
 30 複数種類の言語のそれぞれで用いられる文字を認識する複数の認識工程と、
 前記画像データに関し、前記複数の認識工程による認識を実行する実行工程と、
 前記認識工程による認識で得られる夫々の認識結果の認識率を獲得する獲得工程と、
 前記獲得工程によって獲得される前記複数の認識率に基づいて、複数の認識結果のいずれか一つを選択する選択工程と、
 前記選択工程によって選択される認識結果を出力する出力工程とを備えることを特徴とする文字認識方法。

【請求項15】 前記選択工程は、前記複数の認識結果のそれぞれの認識率を比較する比較工程を更に備え、
 前記比較工程の比較の結果、認識率が一番高い認識結果を選択することを特徴とする請求項14に記載の文字認識方法。

【請求項16】 前記選択工程は、前記認識結果のそれぞれの認識率を所定の閾値とで比較する比較工程とを更に備え、
 40 前記比較工程の比較の結果、前記所定の閾値より大きい認識結果を選択することを特徴とする請求項14に記載

の文字認識方法。

【請求項17】 前記原稿を光学的に読み取る読取工程と前記読取工程によって読み取られる画像データの属性に基づいて、前記画像データを属性毎の領域に分離する分離工程とを更に備えることを特徴とする請求項14に記載の文字認識方法。

【請求項18】 前記認識工程は、前記分離手段によって分離される領域の属性が文字である領域を画像データとして認識することを特徴とする請求項17に記載の文字認識方法。

【請求項19】 前記領域を所定の単位で分割する分割工程とを更に備えることを特徴とする請求項17に記載の文字認識方法。

【請求項20】 前記所定の単位は、少なくとも行、文字を含むことを特徴とする請求項19に記載の文字認識方法。

【請求項21】 原稿を読み取って得られる画像データの文字認識を行う文字認識方法であって、複数種類の言語のそれぞれで用いられる文字を認識する複数の認識工程と、前記複数の認識工程の一つを選択し、選択される認識工程による前記画像データの認識を実行する実行工程と、前記認識工程の認識によって得られる認識結果の認識率を獲得する獲得工程と、前記獲得工程によって獲得された前記認識率と所定の閾値とを比較する比較工程と、前記比較工程の比較の結果、前記認識率が前記所定の閾値よりも大きい場合、前記認識結果を出力する出力工程と、

前記出力工程による出力結果が得られるまで、前記認識工程の選択を順次切り換えて前記実行工程の実行がなされるよう制御する制御工程とを備えることを特徴とする文字認識方法。

【請求項22】 前記認識工程の使用順序を設定する設定工程を更に備え、

前記切り換えは、前記設定工程で設定される順序に従って切り換えることを特徴とする請求項17に記載の文字認識方法。

【請求項23】 前記原稿を光学的に読み取る読取工程と前記読取工程によって読み取られる画像データの属性に基づいて、前記画像データを属性毎の領域に分離する分離工程とを更に備えることを特徴とする請求項21に記載の文字認識方法。

【請求項24】 前記認識工程は、前記分離工程によって分離される領域の属性が文字である領域を画像データとして認識することを特徴とする請求項23に記載の文字認識方法。

【請求項25】 前記領域を所定の単位で分割する分割工程とを更に備えることを特徴とする請求項23に記載の文字認識方法。

【請求項26】 前記所定の単位は、少なくとも行、文字を含むことを特徴とする請求項25に記載の文字認識方法。

【請求項27】 文字認識処理のプログラムが格納されたコンピュータ可読メモリであって、複数種類の言語のそれぞれで用いられる文字を認識する複数の認識工程の手順コードと、

前記画像データに関し、前記複数の認識工程による認識を実行する実行工程の手順コードと、

10 前記認識工程による認識で得られる夫々の認識結果の認識率を獲得する獲得工程の手順コードと、

前記獲得工程によって獲得される前記複数の認識率に基づいて、複数の認識結果のいずれか一つを選択する選択工程の手順コードと、

前記選択工程によって選択される認識結果を出力する出力工程の手順コードとを備えることを特徴とするコンピュータ可読メモリ。

【請求項28】 文字認識処理のプログラムが格納されたコンピュータ可読メモリであって、

20 複数種類の言語のそれぞれで用いられる文字を認識する複数の認識工程の手順コードと、

前記複数の認識工程の一つを選択し、選択される認識工程による前記画像データの認識を実行する実行工程の手順コードと、

前記認識工程の認識によって得られる認識結果の認識率を獲得する獲得工程の手順コードと、

前記獲得工程によって獲得された前記認識率と所定の閾値とを比較する比較工程の手順コードと、

30 前記比較工程の比較の結果、前記認識率が前記所定の閾値よりも大きい場合、前記認識結果を出力する出力工程の手順コードと、

前記出力工程による出力結果が得られるまで、前記認識工程の選択を順次切り換えて前記実行工程の実行がなされるよう制御する制御工程の手順コードとを備えることを特徴とするコンピュータ可読メモリ。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、原稿を読み取って得られる画像データの文字認識を行う文字認識装置及びその方法に関するものである。

【0002】

【従来の技術】情報が溢れる現代社会において情報管理、検索が容易になる情報の電子化が早急に望まれている。情報の電子化には、スキャナ等の読取装置で読み取る画像の中で、特に画像が文字である場合、その文字を認識し文字コードに変換するOCR（光学的文字認識）が必須であり、その精度はどんどん向上してきている。

【0003】OCRを備える読取装置において、その装置単体で言語が異なる複数種類の文字の認識を行う場合、各言語の特性の違いのため、各言語を精度良くに認

識することはできなかった。例えば、日本語専用のOCRで、アルファベットを認識する場合、特にアルファベットの小文字はその特性が日本語と著しく異なるため、認識することができなかった。

【0004】そのため、1つの読取装置で、複数種類の言語を認識するために、言語別の認識アルゴリズムをそれぞれ用意し、言語毎にその言語に対応する認識アルゴリズムをユーザが入力装置等で切り換えて用いることで精度良い文字認識を行っていた。また、認識アルゴリズムは同じでも、予め装置に各文字の辞書を記憶させ、認識を行う度に認識させる文字に対応する言語の辞書をユーザがオペレーションパネル等の入力装置より切り替えることで、各言語の文字認識を行っていた。更には、各辞書の切り替えを行うための制御も必要であった。

【0005】

【発明が解決しようとする課題】しかしながら、ユーザがオペレーションパネル等の入力装置を使用して各言語の辞書を切り換える指示を行いながら複数種類の言語の文字認識を行う方法は、ユーザの手間が多くなり処理速度を低下させるという問題点があった。また、読取原稿が複数枚ある場合、その原稿の差し替え動作を削減するためにADF（オートドキュメントフィーダー）を用いて文字認識を行っていた。その場合、複数の読取原稿の中に英語の原稿、日本語の原稿が入り混じっていると、読取原稿を一枚読み込む毎に、ユーザの指示を行わなければならない、ADFの利点を損なうばかりか、結果的に処理速度を低下させてしまうという問題点があった。

【0006】本発明は上記の問題点を鑑みてなされたものであり、複数言語のそれぞれに用いられる文字の文字認識を可能とし、かつ文字認識の処理速度を向上する文字認識装置及びその方法を提供することを目的としている。

【0007】

【発明を解決するための手段】上記の目的を達成するための本発明による文字認識装置は以下の構成を備える。即ち、原稿を読み取って得られる画像データの文字認識を行う文字認識装置であって、複数種類の言語のそれぞれで用いられる文字を認識する複数の認識手段と、前記画像データに関し、前記複数の認識手段による認識を実行する実行手段と、前記認識手段による認識で得られる夫々の認識結果の認識率を獲得する獲得手段と、前記獲得手段によって獲得される前記複数の認識率に基づいて、複数の認識結果のいずれか一つを選択する選択手段と、前記選択手段によって選択される認識結果を出力する出力手段とを備える。

【0008】また、好ましくは、前記選択手段は、前記複数の認識結果のそれぞれの認識率を比較する比較手段を更に備え、前記比較手段の比較の結果、認識率が一番高い認識結果を選択する。一番高い認識率の認識結果を選択することで、精度の良い文字認識を行うことができ

るからである。

【0009】また、好ましくは、前記選択手段は、前記認識結果のそれぞれの認識率を所定の閾値とで比較する比較手段とを更に備え、前記比較手段の比較の結果、前記所定の閾値より大きい認識結果を選択する。また、好ましくは、前記原稿を光学的に読み取る読取手段と前記読取手段によって読み取られる画像データの属性に基づいて、前記画像データを属性毎の領域に分離する分離手段とを更に備える。

10 【0010】また、好ましくは、前記認識手段は、前記分離手段によって分離される領域の属性が文字である領域を画像データとして認識する。また、好ましくは、前記領域を所定の単位で分割する分割手段とを更に備える。また、好ましくは、前記所定の単位は、少なくとも行、文字を含む。

【0011】上記の目的を達成するための本発明による文字認識装置は以下の他の構成を備える。即ち、原稿を読み取って得られる画像データの文字認識を行う文字認識装置であって、複数種類の言語のそれぞれで用いられる文字を認識する複数の認識手段と、前記複数の認識手段の一つを選択し、選択される認識手段による前記画像データの認識を実行する実行手段と、前記認識手段の認識によって得られる認識結果の認識率を獲得する獲得手段と、前記獲得手段によって獲得された前記認識率と所定の閾値とを比較する比較手段と、前記比較手段の比較の結果、前記認識率が前記所定の閾値よりも大きい場合、前記認識結果を出力する出力手段と、前記出力手段による出力結果が得られるまで、前記認識手段の選択を順次切り換えて前記実行手段の実行がなされるよう制御する制御手段とを備える。

【0012】また、好ましくは、前記認識手段の使用順序を設定する設定手段を更に備え、前記切り換えは、前記設定手段で設定される順序に従って切り換える。ユーザの用途に応じて、順序を切り換えることで、処理速度を向上することができるからである。また、好ましくは、前記原稿を光学的に読み取る読取手段と前記読取手段によって読み取られる画像データの属性に基づいて、前記画像データを属性毎の領域に分離する分離手段とを更に備える。

40 【0013】また、好ましくは、前記認識手段は、前記分離手段によって分離される領域の属性が文字である領域を画像データとして認識する。また、好ましくは、前記領域を所定の単位で分割する分割手段とを更に備える。また、好ましくは、前記所定の単位は、少なくとも行、文字を含む。

【0014】上記の目的を達成するための本発明による文字認識方法は以下の構成を備える。即ち、原稿を読み取って得られる画像データの文字認識を行う文字認識方法であって、複数種類の言語のそれぞれで用いられる文字を認識する複数の認識工程と、前記画像データに関

し、前記複数の認識工程による認識を実行する実行工程と、前記認識工程による認識で得られる夫々の認識結果の認識率を獲得する獲得工程と、前記獲得工程によって獲得される前記複数の認識率に基づいて、複数の認識結果のいずれか一つを選択する選択工程と、前記選択工程によって選択される認識結果を出力する出力工程とを備えることを特徴とする文字認識方法。

【0015】上記の目的を達成するための本発明による文字認識方法を以下の他の構成を備える。即ち、原稿を読み取って得られる画像データの文字認識を行う文字認識方法であって、複数種類の言語のそれぞれで用いられる文字を認識する複数の認識工程と、前記複数の認識工程の一つを選択し、選択される認識工程による前記画像データの認識を実行する実行工程と、前記認識工程の認識によって得られる認識結果の認識率を獲得する獲得工程と、前記獲得工程によって獲得された前記認識率と所定の閾値とを比較する比較工程と、前記比較工程の比較の結果、前記認識率が前記所定の閾値よりも大きい場合、前記認識結果を出力する出力工程と、前記出力工程による出力結果が得られるまで、前記認識工程の選択を順次切り換えて前記実行工程の実行がなされるよう制御する制御工程とを備える。

【0016】上記の目的を達成するための本発明によるコンピュータ可読メモリは以下の構成を備える。即ち、文字認識処理のプログラムが格納されたコンピュータ可読メモリであって、複数種類の言語のそれぞれで用いられる文字を認識する複数の認識工程の手順コードと、前記画像データに関し、前記複数の認識工程による認識を実行する実行工程の手順コードと、前記認識工程による認識で得られる夫々の認識結果の認識率を獲得する獲得工程の手順コードと、前記獲得工程によって獲得される前記複数の認識率に基づいて、複数の認識結果のいずれか一つを選択する選択工程の手順コードと、前記選択工程によって選択される認識結果を出力する出力工程の手順コードとを備える。

【0017】上記の目的を達成するための本発明によるコンピュータ可読メモリは以下の他の構成を備える。即ち、文字認識処理のプログラムが格納されたコンピュータ可読メモリであって、複数種類の言語のそれぞれで用いられる文字を認識する複数の認識工程の手順コードと、前記複数の認識工程の一つを選択し、選択される認識工程による前記画像データの認識を実行する実行工程の手順コードと、前記認識工程の認識によって得られる認識結果の認識率を獲得する獲得工程の手順コードと、前記獲得工程によって獲得された前記認識率と所定の閾値とを比較する比較工程の手順コードと、前記比較工程の比較の結果、前記認識率が前記所定の閾値よりも大きい場合、前記認識結果を出力する出力工程の手順コードと、前記出力工程による出力結果が得られるまで、前記認識工程の選択を順次切り換えて前記実行工程の実行が

なされるよう制御する制御工程の手順コードとを備える。

【0018】

【発明の実施の形態】以下、図面を参照して本発明の好適な実施の形態を詳細に説明する。

＜実施の形態1＞図1は実施の形態1の文字認識装置の構成を示すブロック図である。同図において、111はCPUであり、ROM112に記憶されているプログラムに従って、RAM113、キーボード114、文字認識部115、表示部118、記録部117、読取部118の各種制御を行う。112はROMであり、キーボード114から入力されるデータの処理や文字認識部115の文字認識を実行するための各種プログラムを格納している。113はRAMであり、各種プログラムやキーボード114から入力されるデータの作業領域及び一時待避領域である。

【0019】114はキーボードであり、後述する文字認識の開始の指示やデータの入力を行う。115は文字認識部であり、後述する図3のフローチャートで説明される処理を行う。116は表示部であり、文字認識部115の認識結果や処理状況を表示する。117は記録部であり、文字認識部の認識結果等をユーザのキーボード114による指示で、記録媒体に記録する。

【0020】118は読取部であり、光学的に画像の読み取りを行う光学式文字読取装置（OCR）からなる読取部である。また、読み取られた画像データは、1画素1ビットの単純2値化された2値画像であるとする。更に、読み取られた画像データは1ビット単位のビットマップ画像としてRAM111に展開され、後述する図3のフローチャートで説明される処理が行われる。

【0021】尚、読取部118によって読み取られた画像データに傾きが生じている場合、文字認識処理を実行する前に、画像データの傾きを補正することによって、原稿が傾いた画像データに対してより良好な結果を得ることができる。傾きの補正方法としては、例えば、抽出した行の傾きを求め、その傾きをなくするような座標変換を実行すれば良い。

【0022】119は、FDD（フロッピーディスクドライブ）であり、FDを装着しデータの読み書きが可能である。また、装着されたFD（不図示）に後述するフローチャートのプログラムを書き込み、該プログラムを本装置のRAM112に読み込むことで処理を実行させることもできる。尚、FDD119のかわりにCD-ROMドライバやHDDを備え、それぞれのドライバに装着あるいは内蔵されるCD-ROMやHDDに上述のプログラムを記憶させて、記憶されたプログラムを読み出すことで処理を実行させることも容易に可能である。

【0023】120はCPUバスであり、文字認識装置の構成要素を相互に接続する。尚、実施の形態1で、読取部118で読み取られる読取原稿は、日本語あるいは

英語あるいは日本語と英語が入り交じった原稿であるとする。次に文字認識部115の詳細な説明を図2を用いて説明する。図2は実施の形態1の文字認識部115の詳細な構成を示すブロック図である。

【0024】101は認識単位切り出し部であり、読取部118で読み取られた画像データの認識を所定の大きさを持つ認識単位毎に行うために、画像データを複数のブロックに分割する。尚、認識単位とは、例えば、読取画像を後述する領域分離により所定の領域内に占める文字の集まりである文字ブロック、あるいは文字ブロック

10を単位で分割した行、あるいは行を文字単位で分割した文字のことである。
【0025】102は日本語の文字認識を行う認識部1であり、認識単位切り出し部101により認識単位毎に分割された画像データの認識を行う。また、認識部1(102)で認識された結果、認識結果1(図2の(b))はセレクタ105に出力される。更に、認識結果1の確からしさを示す認識自信度C1(図2の(a))は判定部104へ出力される。

【0026】103は英語の文字認識を行う認識部2であり、認識単位切り出し部101により認識単位毎に分割された画像データの認識を行う。また、認識部2(103)で認識された結果、認識結果2(図2の(c))はセレクタ105に出力される。104は判定部であり、認識部1(102)により出力された認識自信度C1(図2の(a))を入力し、その値に基づいて1ビットの判定信号をセレクタ105へ出力する。尚、判定部104で出力する判定信号は、例えば、認識自信度C1と所定の閾値T1を比較し、認識自信度C1が閾値T1よりも大きい場合に「1」を出力し、小さい場合に「0」を出力する。また、自信度Cの上位1ビットを取ってそれを判定信号としてもよい。

【0027】105はセレクタであり、認識部1(102)により認識された認識結果1と認識部2(103)により認識された認識結果2及び判定部104により判定された判定信号が入力される。判定部104より入力される判定信号に基づいて、認識結果1または認識結果2のどちらかが選択されて、それが認識結果として出力される。

【0028】次に実施の形態1における処理について、図3のフローチャートを用いて説明する。図3は実施の形態1の処理フローを示すフローチャートである。尚、実施の形態1では、認識単位切り出し部101に分割される認識単位を「行」とする。

【0029】ステップS1101で、読取部118より読み取られた画像データは、領域分離処理により以下に示すようなブロックに分割される。例えば、読取部118より読み取られた画像データが、図4の(a)に示されるようなものであったとする。この画像データを、画像データの種類(例えば、タイトル、テキスト、図、

絵)に応じて、図4の(b)に示されるように種類分けを行う処理が領域分離である。ここで、領域分離処理の処理フローを図5のフローチャートを用いて説明する。

【0030】ステップS301で、読み取った画像データを $m \times n$ 画素単位の領域毎に対しウインドリング処理を行い、1画素でも黒ドットがある領域には、その領域内の画像データが連結する程度に解像度を低くする解像度変換を行う。例えば、図4の(a)に示された画像データに対し解像度変換を実行すると、図6に示されるような画像となる。

【0031】尚、明らかにノイズとわかる孤立ドットなどによって得られる領域は、例えば、ウインドリング処理後、パターンマッチング等で消去する。ステップS302で、解像度変換された画像に対し、各画像の輪郭線の特徴を把握するための輪郭線追跡を行う。この処理によって、輪郭線の特徴が細長いパターンである場合はテキストあるいはタイトルと判定され、それ以外の場合は図形あるいは絵として区別することができる。

【0032】ステップS303で、同一種類の画像が隣接する場合は、隣接する画像同士をつなげる連結処理を行う。この処理によって、図4の(a)の画像データは、図4の(b)のような領域分離処理が完了する。

尚、この領域分離処理完了時に、図7の(a)に示すような画像データの種類等を示すデータ構造体と、図7の(b)に示すような領域分離されたブロックの配置関係を示すブロックデータが得られる。また、このデータ構造体とブロックデータに従って、文字認識の認識する順番等の処理に関わる情報を適宜獲得する。

【0033】図7の(a)のデータ構造体と(b)のブロックデータについて説明する。まず、図7の(a)で示されるデータ構造体は、領域分離によって分割されるブロック(struct BLOCK)の種類(図のshort type)と位置関係(図のshort startx, starty)と各ブロックの幅、高さ(図のshort width, height)と各ブロックの順番を決定するアドレス(struct BLOCK *next_addresses)の定義を行っている。

【0034】図ではブロックの種類として、0=TITLE(タイトル)、1=TEXT(テキスト)、2=FIGURE(図)、3=PICTURE(絵)と定義する。各ブロックに定義される数字が、領域分離によって決定されるブロックの種類としてブロックデータで用いられる。また、各ブロックの位置関係は、図4の(b)に示されるような、x、y座標に対する位置関係であり、startxが各ブロックのx座標、startyが各ブロックのy座標である。尚、各ブロックの位置は、各ブロックの所定の位置関係にある角(図4の(b)の●印で示される角)を各ブロックの位置としている。更に、各ブロックの幅、高さは、各ブロックのx方向の長さ(widthで定義される幅)であり、y方向の長さ(heightで定義される高さ)である。

【0035】図5の(b)に示されるように、図4の(a)で示される画像データが領域分離により5つのブロック(図4の(b)参照)に分割されるので、各ブロックの処理の順番(アドレス)としてBLK[0]~BLK[4]が与えられている。そして、各ブロックでは、図7の(a)で説明されたブロックの種類と座標とその幅、高さ、更に次に処理すべきブロックのアドレスが示される。

【0036】例えば、BLK[0]で説明すると、BLK[0]で示されるブロックは図4の(b)のテキストであったとすると、ブロックの種類であるtypeが「0」となる。また、位置としてx座標であるstartxが「500」、y座標であるstartyが「300」となる。更に、幅であるwidthが「1500」、高さであるheightが「250」となる。更にまた、次に処理すべきブロックのアドレスであるnext_addressが「&BLK[1]」となる。同様に、BLK[1]~BLK[4]についても説明される。尚、最後に処理されるブロック(図ではBLK[4])については、次に処理すべきブロックが存在しないので、次に処理すべきブロックのアドレスであるnext_addressは「NULL」となる。

【0037】図3のフローチャートの説明に戻る。ステップ1102で、ステップ1101で領域分離によって分割された各ブロックについて以下に説明されるステップS1103以降の処理が完了しているか否かを判定する。つまり、各ブロックの順番に従って処理が行われ、next_addressが「NULL」であるブロックの処理が完了するとすべてのブロックにおける処理が完了したことになる。

【0038】ステップS1103で、現在処理しようとするブロックが文字を含んでいる(typeが「0」または「1」)か否かを判定する。文字を含んでいる場合(ステップS1103でYES)、ステップS1104に進む。文字を含んでいない場合(ステップS1103でNO)、ステップS1102に戻る。ステップS1104で、ブロック内の画像データ(文字)をy軸方向に射影を行い、文字の存在するy軸(行)の抽出を行う。y軸方向への射影の例として、図4の(b)のテキスト2に対して行くと、図8に示すように文字がある部分(行)に対してはy'上に射影が現れる。具体的なy軸方向への射影方法について図9のフローチャートを用いて説明する。

【0039】図9は実施の形態1のy軸方向への射影方法の処理フローを示すフローチャートである。尚、各ブロックに存在する文字は、1ビットのビットマップ画像であるので、y軸方向への射影は、1ビット単位でブロック内の画像について「黒」であるか「白」を判定しながら検索し、その判定と検索結果に基づいて行う。また、y軸方向への射影の処理開始時点では、ステップS901~ステップS905で説明される値の初期化が行

われる。

【0040】ステップS901で、文字の画像データを含むブロックにおいて、そのブロックに占める行数を示すnlineをリセットする。ステップS902で、行の始まりを検索しているのか、あるいは行の終わりを検索しているのかを区別するフラグflagをリセットする。ステップS903で、y方向の検索回数をカウントするカウンタjをリセットする。ステップS904で、x方向の検索回数をカウントするカウンタiをリセットする。ステップS905で、カウンタiが示す値に位置する画素が黒であるか否かを示すフラグkuroをリセットする。

【0041】ステップS906で、座標(startx+i, starty+j)に位置する画素が黒であるか否かを判定する。画素が黒である場合(ステップS906でNO)、ステップS907に進む。画素が黒でない場合、つまり、白である場合(ステップS906でYES)、ステップS911に進む。ステップS911で、startx+iがブロックの幅widthより大きいのか否かを判定する。大きい場合(ステップS911でNO)、ステップS912に進む。小さい場合(ステップS911でYES)、ステップS916に進む。ステップS916で、iの値を+1インクリメントして、ステップS906に戻る。

【0042】一方、ステップS907で、kuroが1をセットする。ステップS908で、flagが「0」であるか否か(行の始まりを探しているモードなのか否か)を判定する。flagが「0」である場合(ステップS908でYES)、ステップS909へ進む。flagが「0」でない場合(ステップS908でNO)、ステップS912へ進む。

【0043】ステップS909で、y座標がstarty+jであるラインを示すline_sy[nline]にjの値を代入する。ステップS910で、flagを「1」に変更する。ステップS912で、kuroが「0」(そのラインに黒画素が無く)かつflagが「1」(行の終わりを探しているモード)であるか否かを判定する。kuroが「0」かつflagが「1」である場合(ステップS912でYES)、ステップS913に進む。kuroが「0」かつflagが「1」でない場合(ステップS912でNO)、ステップS917に進む。

【0044】ステップS913で、射影により抽出されたnline番目の行の高さを示すline_h[nline]に、現在の処理ラインであるjからその行のスタート位置line_sy[nline]を代入する。ステップS914で、行抽出か否かを示すflagを行の始まりを探すモードに戻し、ステップS915で、行数を示すnlineを+1インクリメントしてステップS917に飛ぶ。一方、kuroが「0」かつflagが「1」でない場合(ステップS912でNO)、ステップS917に飛び、iの値を+1インクリメントし、ステップS904に戻る。

【0045】以上の図9の処理により文字画像から行

数、各行のスタート点および高さを抽出することができる。尚、図のフローチャートで説明される行の抽出方法は、ブロック内の画像データ（文字）をY軸方向に射影を行い、文字の存在するY軸（行）の抽出を行っていたが、これに限らない。例えば、図5のフローチャートのステップS302の輪郭線追跡を行なったあと、輪郭線追跡により抽出されたオブジェクトのうち細長いものを文字と判定するが、ステップS303で、同一グループ結合を行わずに行として領域分離を実行する。この処理を図4の(a)の画像データに対して実行すると図10のようになり、領域分離の時点で行が抽出される。但し、解像度変換の際に抽出される行の精度が低いため、その精度を考慮した所定の閾値T1の値を設定しなければならない。

【0046】図3のフローチャートの説明に戻る。ステップS1105で、ステップS1104で抽出した各行の認識処理が終了した否かを判定する。全ての行の認識処理が終了した場合（ステップS1105でNO）、ステップS1102に進む。終了していない場合（ステップS1105でYES）、ステップS1106に進む。

【0047】ステップS1106で、認識部1で、行単位の文字認識を行い、認識結果1と認識自信度C1を求める。同時に、認識部2で、行単位の文字認識を行い、認識結果2を求める。ステップS1108で、認識自信度C1が所定の閾値T1より大きいかなかを判定する。認識自信度C1が所定の閾値T1より大きい場合（ステップS1108でYES）、ステップS1109に進む。認識自信度C1が所定の閾値T1より小さい場合（ステップS1108でNO）、ステップS1110に進む。ステップS1109で、認識結果1を認識結果として選択する。ステップS1110で、認識結果2を選択結果として選択する。以上、ステップS1106～ステップS1110の認識処理を全ての行に対して行うことで、ブロック内に含まれる全ての行の文字認識が完了する。

【0048】以上説明したように、実施の形態1によれば、日本語用の文字認識部1と英語用の文字認識部2を設け、読取原稿の認識処理を並列的にそれぞれの認識部で行う。そして、認識部1によって求まる認識自信度C1に基づいて、認識結果を決定することで、ユーザが読取原稿を読取部118を読み取らせる前に、読取原稿が英語であるか日本語であるかを確認することなく、文字認識を行うことができる。

【0049】尚、読取部118で、テキストあるいはタイトルである画像データの読取時に、その画像データの方向を検出することによって、読取原稿の方向に依存しない読取部118で構成することも容易に可能である。画像データの方向の検出方法としては、例えば、領域分離によって抽出したテキストあるいはタイトルであるブロックからある行（画像データ）を抽出し、抽出した行

（画像データ）を0°回転、90°回転、180°回転、270°回転を行った行（画像データ）を獲得する。そして、それぞれ獲得された画像データに対し、読取部118によって読み取りを行う。その結果、読み取られた画像データに対する各回転の認識度を得ることにより、テキストあるいはタイトルである画像データの方向を検出し、検出された方向に基づいて読取原稿の回転を行えば良い。

【0050】尚、実施の形態1の図3のフローチャートの処理を、以下、図11のフローチャートで説明される処理で行っても良い。尚、図11において、ステップS1301からステップS1305の処理は、図11のステップS1101からステップS1105の処理と同じであるので、その説明は省略する。

【0051】ステップS1306で、認識部1で、文字認識し認識結果1及び認識自信度C1を求める。ステップS1307で、認識自信度C1が所定の閾値T1より大きいかなかを判定する。認識自信度C1が所定の閾値T1より大きい場合（ステップS1307でYES）、ステップS1308に進む。認識自信度C1が所定の閾値T1より小さい場合（ステップS1307でNO）、ステップS1308に進む。ステップS1308で、認識結果1を認識結果として選択する。

【0052】一方、ステップS1309で、認識部2で、文字認識し認識結果2及び認識自信度C2を求める。ステップS1310で、認識結果2を認識結果として選択する。以上、ステップS1306～ステップS1310の認識処理を全ての行に対して行うことで、ブロック内に含まれる全ての行の文字認識が完了する。図11で説明される処理を実行することにより、認識部2における文字認識はステップS1307で、認識自信度C1が所定の閾値T1未満である場合だけに限って行われる。そのため、認識結果1による認識処理だけで文字認識が完了する場合は、認識部2による処理を行わずに済むので、処理の高速化を図ることができる。

【0053】尚、実施の形態1では、認識部1における認識自信度C1を出力していたが、これに限らない。例えば、図12に示すように認識部2により、認識自信度C2（図の(d)）を出力する構成にしても良い。この場合、判定部104において、例えば、認識自信度C1と認識自信度C2を比較し認識自信度の高い方を、認識結果として出力するように判定信号をセレクト105へ出力する。その結果、認識結果の良い方が認識結果としてセレクト105で選択される。また、認識部1と認識部2の認識自信度のレンジまたは位置付けが違う場合は、両者のレンジまたは位置付けを同等にする変換部（不図示）を判定部104に備え、その変換部の処理を経た後、各認識部による認識結果の比較を行えば良い。尚、実施の形態1では、認識部において、タイトルあるいはテキストと判定されたブロックの文字認識単位を行

単位とし、そのブロックに含まれる各行毎に文字認識を行ったが、これに限らない。例えば、タイトルあるいはテキストと判定されたブロックの1行目を抽出し、その1行目に対し認識部1と認識部2による文字認識処理を行う。そして、認識結果として認識部1あるいは認識部2のどちらかが選択されたら、そのブロック内の画像データ残りのすべてに対して、その選択された認識部を用いて文字認識を行ってもよい。

【0054】また、タイトルあるいはテキストと判定されたブロックの画像データをy軸方向への射影により行の抽出を行い、x軸方向への射影により文字の抽出を行なう。そして、行と文字に分割されたブロックの中から任意のm文字を抽出し、その抽出したm文字に対し認識部1と認識部2による文字認識処理を行う。そして、認識結果として認識部1あるいは認識部2のどちらかが選択されたら、そのブロック内の画像データ残りのすべてに対して、その選択された認識部を用いて文字認識を行ってもよい。

【0055】このような構成にすることで、文字認識処理にかかる処理時間を減少させることができる。尚、実施の形態1では、図2の認識単位切り出し部101による領域分離処理を施す画像データは、1画素1ビットの画像であったが、これに限らない。例えば、1画素8ビットの多値画像に関して領域分離を行うこともできる。この場合の領域分離処理は、画像データに対し微分フィルタをかけることで、画像データの高周波成分を抽出し得られた結果よりテキストデータ、イメージデータに領域分離を行う。また、実施の形態1のように、英語と日本語の判別する必要がある場合は、1画素8ビットの多値画像データを一定の閾値を基準に2値化データにしてから領域分離処理を行う。

【0056】＜実施の形態2＞実施の形態1では、日本語用と英語用の認識部をそれぞれ備え、日本語と英語の文字認識を行う構成であったが、複数の言語に対してそれぞれの認識部を拡張して備えることで、日本語、英語に限らない複数の言語の文字認識を同時に行うことができる。

【0057】例えば、図13に示すように、 n (n は正の整数)個の言語それぞれに対する認識部1(102)、認識部2(103)、…、認識部 n (106)を用意する。そして、各認識部1(102)～認識部 n (106)より、認識結果1、認識結果2、…、認識結果 n をセレクタ105へ出力する。また、認識自信度 C_1 、認識自信度 C_2 、…、認識自信度 C_n を判定部104へ出力する。判定部104は、入力された認識自信度 $C_1 \sim C_n$ を比較し、その結果に基づいて $\log_2(n)$ ビットの判定信号をセレクタ105へ出力する。そして、セレクタ105は判定信号に従って、認識結果として認識結果1～ n の中から一つ選択される。

【0058】例えば、3個の認識部で構成される場合の

文字認識処理を図14のフローチャートを用いて説明する。図14は実施の形態2の処理フローを示すフローチャートである。ステップS1501で、認識単位切り出し部101で画像データの認識単位毎による領域分離を行う。ステップS1502で、ステップS1501で領域分離によって分割された各ブロックについて文字認識処理が完了しているか否かを判定する。各ブロックについて文字認識処理が完了していない場合(ステップS1502でNO)、処理を終了する。各ブロックについて文字認識が完了していない場合(ステップS1502でYES)、ステップS1503に進む。

【0059】ステップS1503で、認識部1で、文字認識し認識結果1及び認識自信度 C_1 を求める。ステップS1504で、認識自信度 C_1 が所定の閾値 T_1 より大きいかな否かを判定する。認識自信度 C_1 が所定の閾値 T_1 より大きい場合(ステップS1504でYES)、ステップS1505に進む。認識自信度 C_1 が所定の閾値 T_1 より小さい場合(ステップS1504でNO)、ステップS1506に進む。ステップS1505で、認識結果1を認識結果として選択する。

【0060】ステップS1506で、認識部2で、文字認識し認識結果2及び認識自信度 C_2 を求める。ステップS1507で、認識自信度 C_2 が所定の閾値 T_2 より大きいかな否かを判定する。認識自信度 C_2 が所定の閾値 T_2 より大きい場合(ステップS1507でYES)、ステップS1508に進む。認識自信度 C_2 が所定の閾値 T_2 より小さい場合(ステップS1507でNO)、ステップS1509に進む。ステップS1508で、認識結果1を認識結果として選択する。

【0061】ステップS1509で、認識部3で、文字認識し認識結果3を求める。ステップS1510で、認識結果3を認識結果として選択する。以上、ステップS1504～ステップS1510の認識処理をすべてのブロックに行うことで、各ブロックの認識処理が完了する。尚、実施の形態2では、認識部を3個としたが、これに限らない。認識を実行させたい複数の言語の数に応じて、いくつでも拡張することができる。例えば、 n 個の認識部から構成される場合、認識部 n の認識結果 C_n と認識自信度 T_n とすると、 $C_n \geq T_n$ の条件を満たすならば認識結果 n を選択し、満たさないならば認識部 $(n+1)$ 以降の認識部による処理を実行し、各認識部において上述の条件を満たす認識部の認識結果を認識結果として選択する。

【0062】また、認識部1～ n の認識させる順番は、予め決まった所定の順序に従うものとするが、ユーザの用途に応じては、キーボード等の入力装置を介してその順番を変更できるものとする。以上説明したように、実施の形態2によれば、複数の言語に対応する認識部を設けることで、同時に複数の言語の文字認識を行うことが可能となる。

【0063】尚、上記FDDの機能若しくは方法の機能によって達成される本発明の目的は、前述の実施の形態のプログラムを記憶させたFD等の記憶媒体によっても達成できる。即ち、上記FDDに、その記憶媒体を装着し、その記憶媒体から読み出したプログラム自体が本発明の新規な機能を達成するからである。このための本発明にかかるプログラムの構造的特徴は図15、図16に示すとおりである。

【0064】実施の形態1の制御を実現するために、図15に示すようにFDには5つのモジュールで構成されている。図に示すように認識モジュール1011、実行モジュール1012、獲得モジュール1013、選択モジュール1014、出力モジュール1015の5つである。この記憶媒体に記憶されるモジュールに従って、図15に示すようなステップS1001～ステップS1005、「認識」、「実行」、「獲得」、「選択」、「出力」の順で処理が実行される。それぞれの記憶媒体に記憶されるモジュールにおいて、認識モジュール1011で実行される「認識」(ステップS1001)処理は、図3のフローチャートのステップS1101～ステップS1105に対応する。

【0065】実行モジュール1012で実行される「実行」(ステップS1002)処理と獲得モジュール1013で実行される「獲得」(ステップS1003)処理は、図3のフローチャートのステップS1106に対応する。選択モジュール1014で実行される「選択」(ステップS1004)処理は、図3のフローチャートのステップS1108に対応する。出力モジュール1014で実行される「出力」(ステップS1005)処理は、図3のフローチャートのステップS1109、ステップS1110に対応する。

【0066】また、実施の形態2の制御を実現するために、図16に示すようにFDには6つのモジュールで構成されている。図に示すように認識モジュール2011、実行モジュール2012、獲得モジュール2013、比較モジュール2014、出力モジュール2015、制御モジュール2016の6つである。この記憶媒体に記憶されるモジュールに従って、図16に示すようなステップS2001～ステップS2006、「認識」、「実行」、「獲得」、「比較」、「出力」、「制御」の順で処理が実行される。それぞれの記憶媒体に記憶されるモジュールにおいて、認識モジュール2011で実行される「認識」(ステップS2001)処理は、図14のフローチャートのステップS1501、ステップS1502に対応する。

【0067】実行モジュール2012で実行される「実行」(ステップS2002)処理と獲得モジュール2013で実行される「獲得」(ステップS2003)処理は、図14のフローチャートのステップS1503に対応する。選択モジュール2014で実行される「選択」

(ステップS2004)処理は、図14のフローチャートのステップS1504に対応する。出力モジュール2015で実行される「出力」(ステップS2005)処理は、図14のフローチャートのステップS1505に対応する。制御モジュール2016で実行される「制御」(ステップS2006)処理は、図14のフローチャートのステップS1506～ステップS1510に対応する。

【0068】また、本発明は、複数の機器から構成されるシステムに適用しても、1つの機器からなる装置に適用しても良い。また、本発明はシステム或は装置にプログラムを供給することによって実施される場合にも適用できることは言うまでもない。この場合、本発明に係るプログラムを格納した記憶媒体が本発明を構成することになる。そして、該記憶媒体からそのプログラムをシステム或は装置に読み出すことによって、そのシステム或は装置が、予め定められた仕方で動作する。

【0069】

【発明の効果】以上の説明からも明らかなように、本発明によれば、複数言語のそれぞれに用いられる文字の文字認識を可能とし、かつ文字認識の処理速度を向上する文字認識装置及びその方法を提供できる。

【図面の簡単な説明】

【図1】実施の形態1の文字認識装置の構成を示すブロック図である。

【図2】実施の形態1の文字認識部115の詳細な構成を示すブロック図である。

【図3】実施の形態1の処理フローを示すフローチャートである。

【図4】実施の形態1の領域分離処理の様子を説明するための図である。

【図5】実施の形態1の領域分離処理の詳細な処理フローを示すフローチャートである。

【図6】実施の形態1の解像度変換の様子を説明するための図である。

【図7】実施の形態1の領域分離処理により作成されるデータ構造体とデータブロックを示す図である。

【図8】実施の形態1のy軸方向への射影によって行が抽出される様子を説明する図である。

【図9】実施の形態1のy軸方向への射影方法の処理フローを示すフローチャートである。

【図10】実施の形態1に他の構成による領域分離処理の結果を示す図である。

【図11】実施の形態1の他の処理フローによるフローチャートである。

【図12】実施の形態1の他の構成による文字認識部115の詳細な構成を示すブロック図である。

【図13】実施の形態1の文字認識部115の詳細な構成を示すブロック図である。

【図14】実施の形態2の処理フローを示すフローチャートである。

ートである。

【図15】実施の形態1の処理フローのプログラムを記憶させたFDDのメモリマップの構造と制御手順を示す図である。

【図16】実施の形態2の処理フローのプログラムを記憶させたFDDのメモリマップの構造と制御手順を示す図である。

【符号の説明】

101 認識単位切り出し部

102 認識部1

103 認識部2

104 判定部

*105 セクタ

106 認識部n

111 CPU

112 ROM

113 RAM

114 キーボード

115 文字認識部

116 表示部

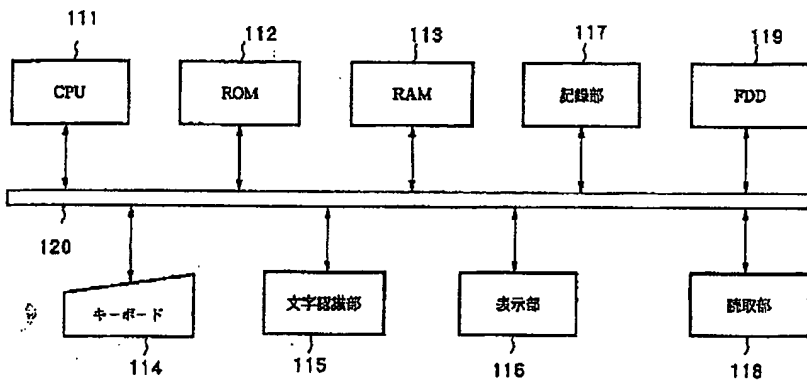
117 記録部

10 118 読取部

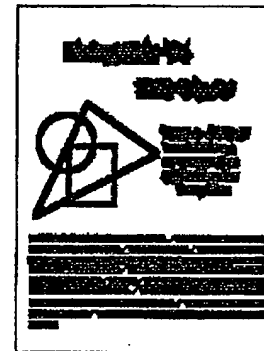
119 FDD

* 120 CPUバス

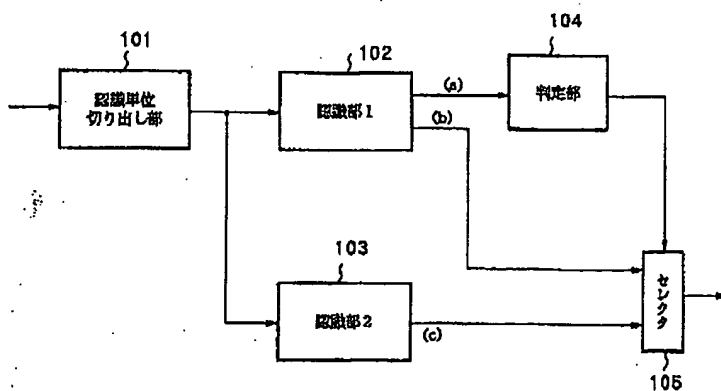
【図1】



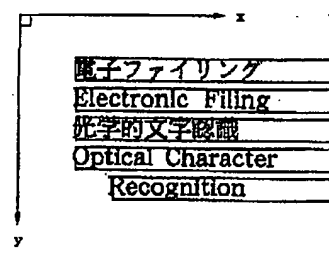
【図6】



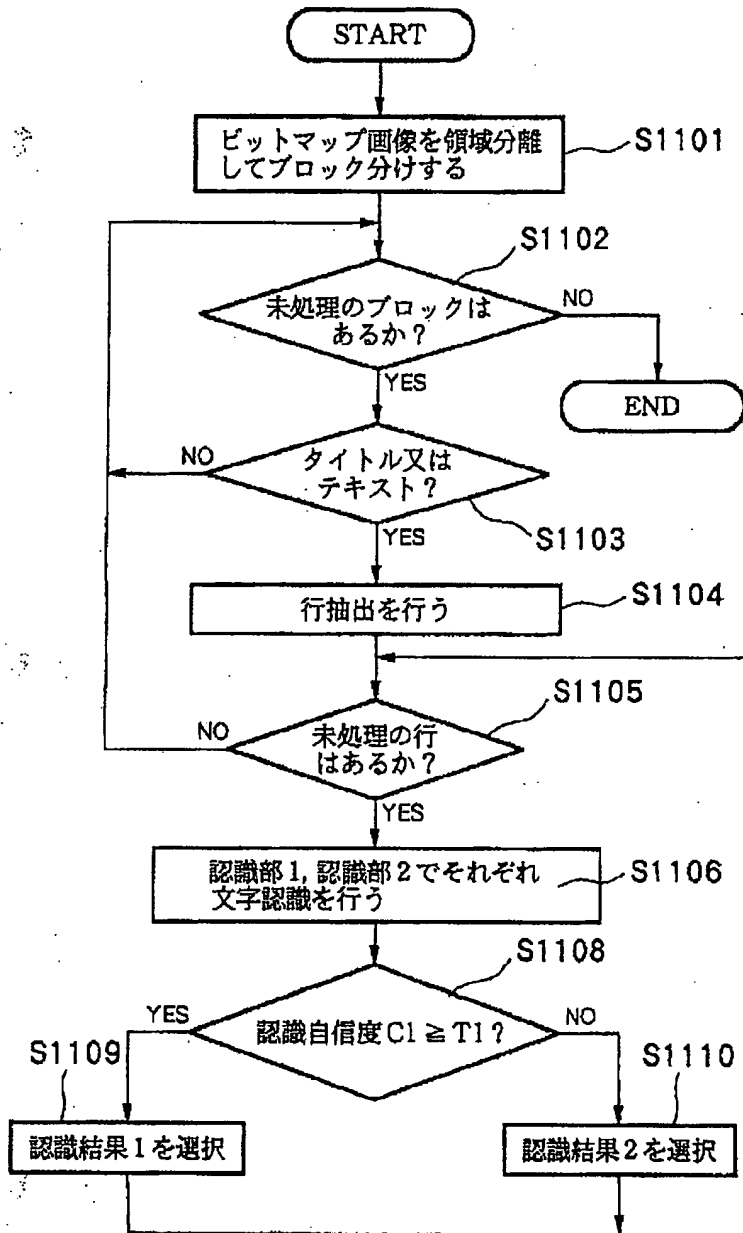
【図2】



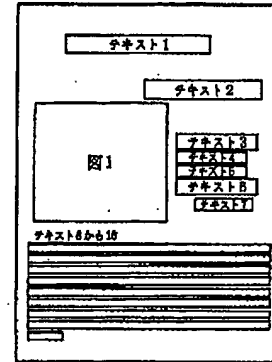
【図8】



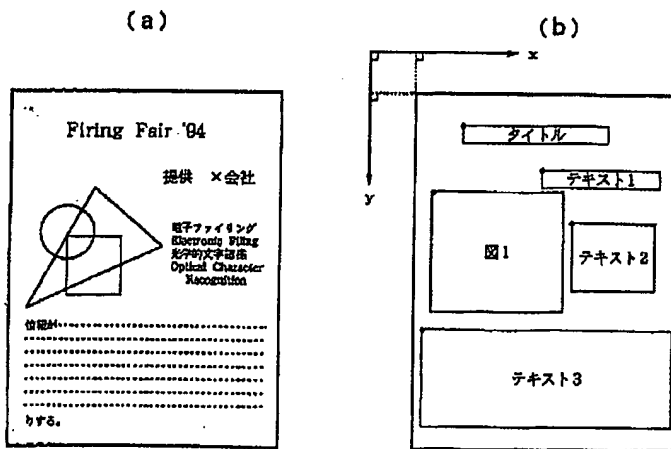
【図3】



【図10】



【図4】



【図7】

(a)

```

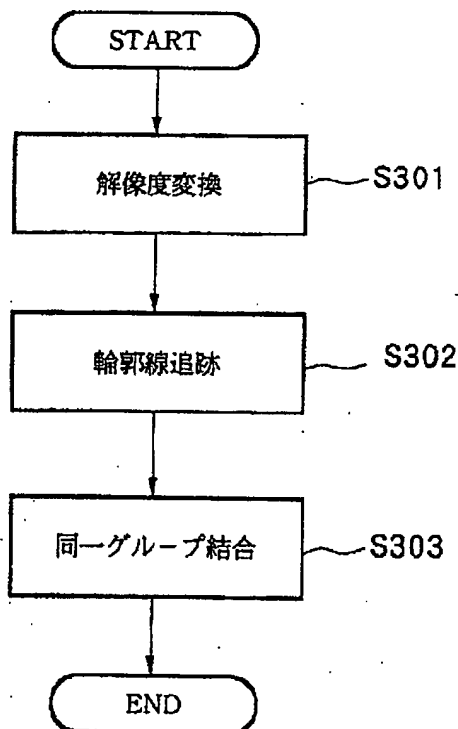
struct BLOCK {
    short type; /* 0=TYITLE, 1=TEXT,
                2=FIGURE, 3=PICTURE */
    short startx, starty;
    short width, height;
    struct BLOCK *next_address;
} BLOCK;
  
```

(b)

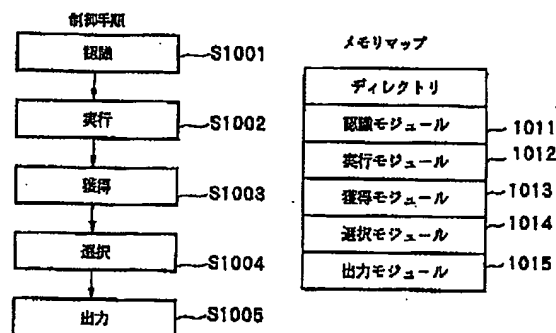
```

BLK [0].type=0;
    .startx=500;
    .starty=300;
    .width=1500;
    .height=250;
    .next_address=&BLK [1].type
BLK [1].type=1;
    .startx=1400;
    .starty=1000;
    .width=1450;
    .height=250;
    .next_address=&BLK [2].type
BLK [2].type=2;
    .startx=160;
    .starty=1300;
    .width=1500;
    .height=1400;
    .next_address=&BLK [3].type
BLK [3].type=1;
    .startx=1700;
    .starty=1500;
    .width=1000;
    .height=800;
    .next_address=&BLK [4].type
BLK [4].type=1;
    .startx=100;
    .starty=2300;
    .width=2800;
    .height=1200;
    .next_address=NULL
  
```

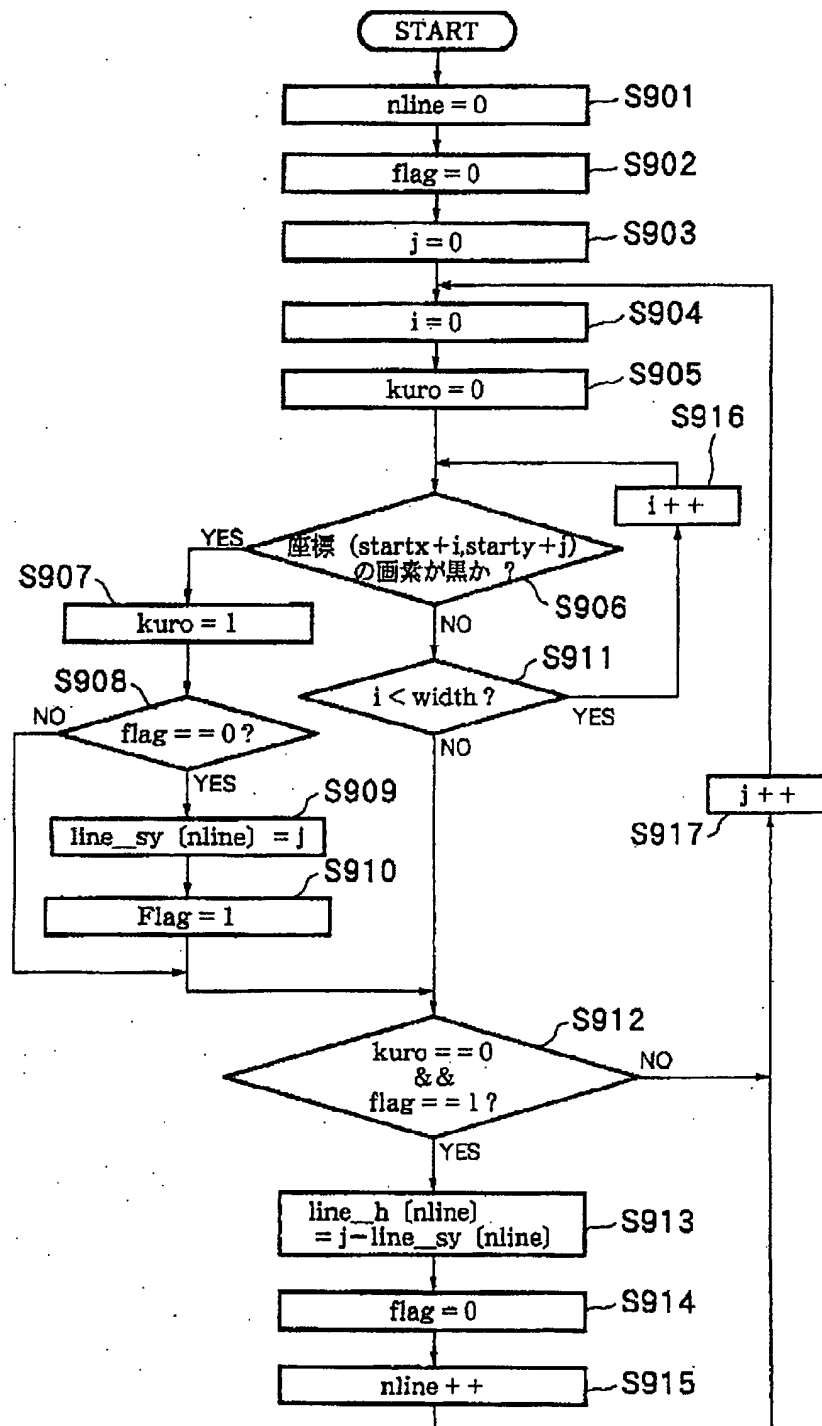
【図5】



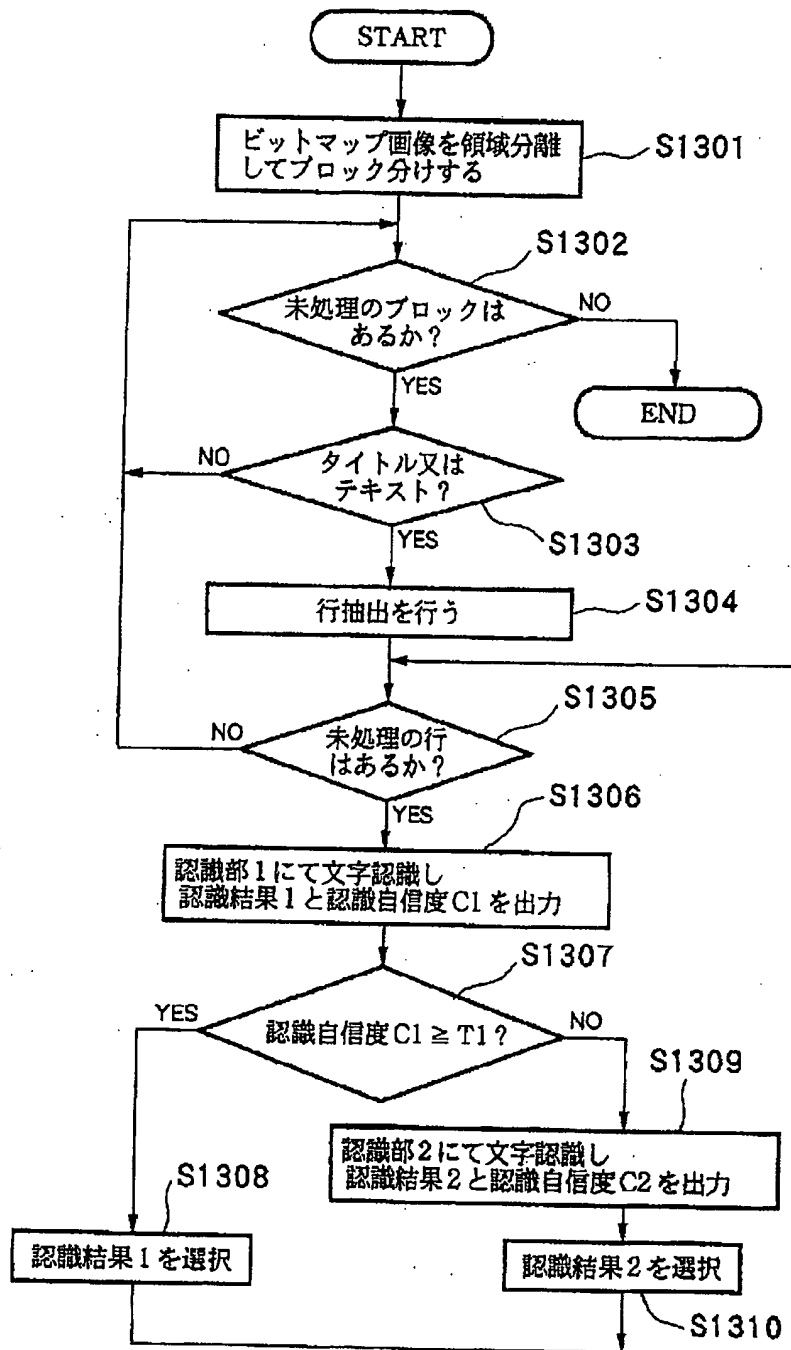
【図15】



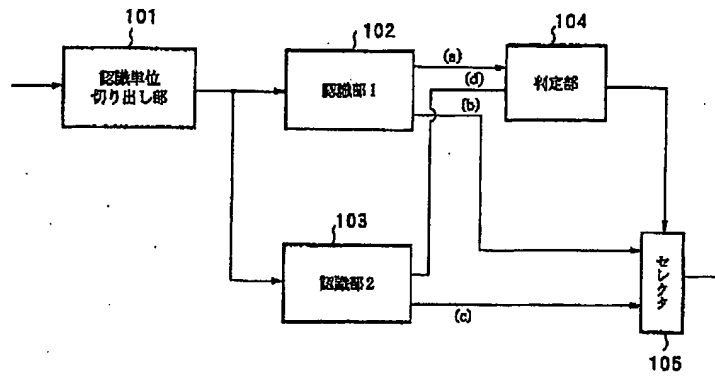
【図9】



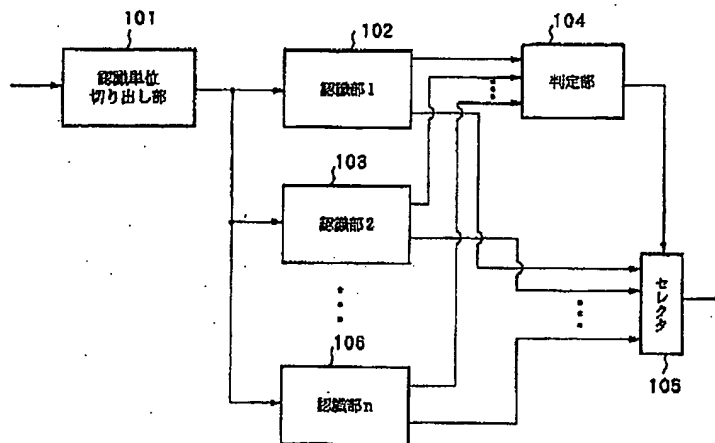
【図11】



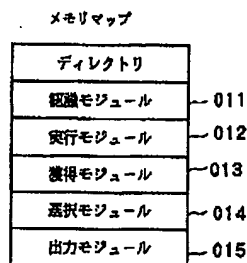
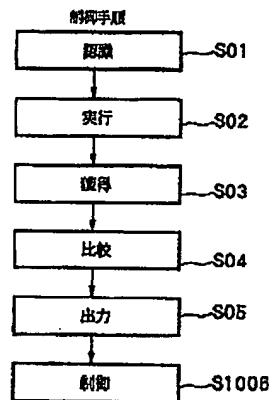
【図12】



【図13】



【図16】



【図14】

